DEFECTIVE WAFER DETECTION USING MUTIPLE CLASSIFIERS

Abderrahmane Boubezoul, Bouchra Ananou, Mustapha Ouladsine, Sebastien Paris[1]
Franck Gasnier, Jacques Pinaton[2]

[1]Laboratory of Sciences of Information's and of System, LSIS UMR 6168, University Paul Cézanne, Aix-Marseille III Av escadrille de Normandie Niemen 13397 Marseille Cedex 20,
BOUBEZOUL.ABDERRAHMANE@LSIS.ORG, TEL. 33 4 91 056028
[2]STMicroelectronics ZI Rousset
13106 Rousset cedex
FRANCK.GASNIER@ST.COM, TEL. 33 4 68 50 45

In a semiconductor factory, Parametric Tests (PT) are performed to check if a particular product is within a certain predefined specifications, and to detect possible process drifts as early as possible. In addition, PT results are used to decide based on statistical analysis whether to accept or to reject a wafer. The method used now is based on univariate statistical techniques which lead to stop many lots unnecessarily.
In this paper, we introduce an automatic wafer classification approach based on combining learning algorithms, which permits the improvement of actual detection performance.
This procedure was successfully validated at PT data provided by STMicroelectronics – Rousset fab, and results show significant reduction of the lot hold rate.

Wafers are subjected to a variety of electrical tests "Parametric Tests", these tests require measuring the electrical parameters of the key devices that form the basic building blocks of all integrated circuits: resistors, capacitors, diodes, transistors, inductors, etc. Measurements are performed on specially designed test structures, typically located in the scribe lines of wafers.
When one or many critical parameters measured on one wafer are out of specificity limits, the lot is declared in "hold" status. "Full Tests" are then applied on these lots in order to detect defective wafers, which will be scraped.
The problem that we have to deal with in this work is the high hold rate compared to the scrap rate (i.e. too many lots are stopped unnecessarily).
The purpose of this work is to propose a decision support system to enable accurate defective wafers detection based on the results of the "Parametric Tests". Such a system would reduce analysis cost by bringing the hold rate closer to the scrap rate.
A historical data base contains "Parametric Tests" results and the decision for each wafer (scrap or not) and (hold or not) for each lot. The dataset is characterized by its high dimension (more than one hundred parameters) and its large portion of missing values and erroneous records.
Before applying our classifiers, a pre-processing is needed. This step involves filling in missing values, smoothing noisy data, identifying and removing outliers.
The data being composed of different parameters with varying order of magnitude, we perform a transformation based on relative residual. It permits to feed into the classifiers data which is solely characterized by the relative deviation from the target. Thus doing, we enhance the predictive capability of our system, making it able to recognize defective wafers belonging to a product which was not utilized in the learning process. This can be of great interest in the context of the initiation of a new product, the quantity of scrap information necessary for the learning will be reduced.
Subsequently, classical standardization and data reduction methods are applied.

Selected algorithms (Generalized Relevance Learning Vector Quantization, Parzen window, K nearest neighbors, Support Vectors Machines) are applied in parallel fashion to our classification problem.
For the final decision we adopt the Weighted Majority Vote approach to combine the decision of the classifiers, and obtain the weights by maximizing the performance of the whole set of classifiers, rather than that of each of them separately (Fig 2). The classification models are not treated equally. Each model is associated with a coefficient (weight), proportional to its classification accuracy.
To assess the performance of the system we distinguish between three categories of wafers in the dataset

- Not Hold – Not Scrap(h_0)
- Hold- Scrap (h_1)
- Hold-Not Scrap (h_2).

To evaluate the performance of the classifiers, we use the cross validation method to build classification models, we divide the dataset in two parts the first (70 % of dataset) is used for training and (30 % of dataset) for testing, and we repeat the training/testing procedures for 100 independent runs.
We obtain a recognition rate of 98.7 % for Train data and 98.32 (±2.43) % for Test data. Classifiers detect (Not Hold – Not Scrap) wafers as good wafers and (Hold – Not Scrap) wafers as good wafers with recognition accuracies 99.9%, 98 % respectively, which leads to significant reduction of the hold rate and all defective wafers were detected.

This methodology based on combining multiple classifiers allow us to reduce the hold rate significantly, these results are promising, and give us an idea about the potential of this approach for resolving such a complex problem.
Finally, this work can be viewed as a step towards building a system for defective wafers detection.

This work is supported by STMicroelectronics – Rousset France and Conseil Général Des Bouches- du-Rhônes

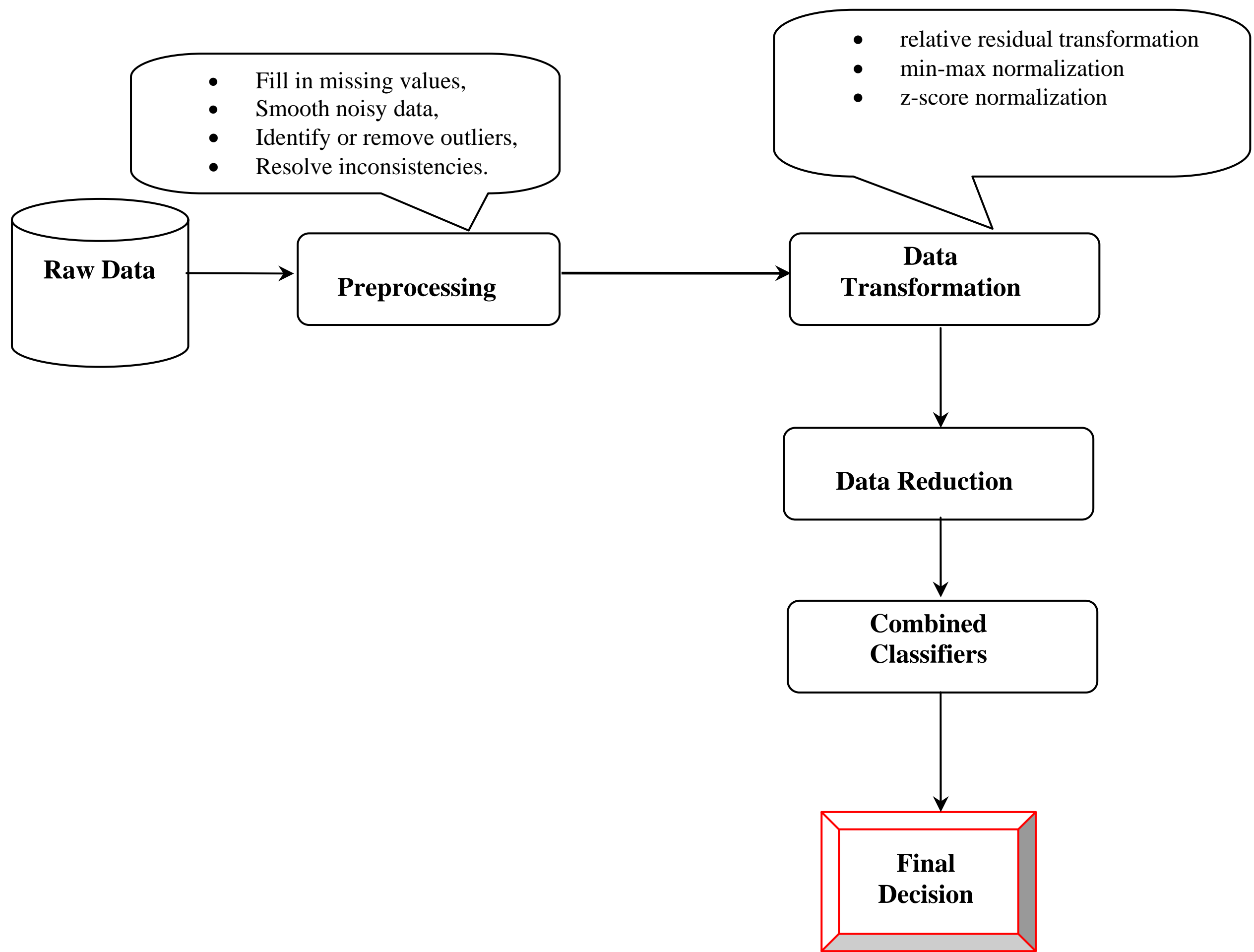

Fig. 1 Preprocessing of data diagram

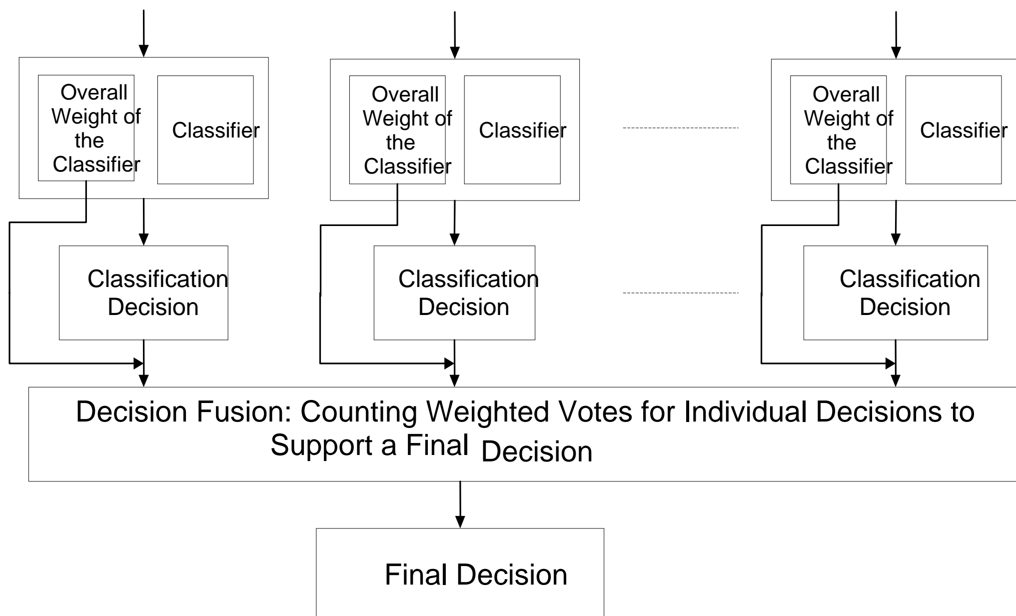

Fig. 2 Weighted Majority Voting diagram